SOURCES OF ERROR AND THE STATISTICAL FORMULATION OF Ms | mb SEISMIC EVENT SCREENING ANALYSIS FOR THE COMPREHENSIVE NUCLEAR-TEST-BAN TREATY (CTBT)

Dale N. Anderson[1], Howard J. Patton[1], Steven R. Taylor[1], Jessie L. Bonner[2], and Neil D. Selby[3]

Los Alamos National Laboratory[1], Weston Geophysical Corporation[2], and AWE Blacknest[3]

Sponsored by the National Nuclear Security Administration

Award Nos. DE-AC52-06NA25396/LA10-ECM-NDD02[1] and FA8718-09-C-0012[2]

ABSTRACT

The Comprehensive Nuclear-Test-Ban Treaty (CTBT), a global ban on nuclear explosions, is currently pre-entry into force. Under the CTBT, a monitoring system of seismic, hydroacoustic, infrasonic and radionuclide sensors operates and data from this system is analyzed by the International Data Centre (IDC). The IDC provides CTBT signatories basic seismic event parameters and a screening analysis indicating whether an event exhibits explosion characteristics (for example, shallow depth). An important component of the screening analysis is a statistical test of the the null hypothesis H0: Explosion Characteristics using empirical measurements of seismic energy (magnitudes). Relative to mb, earthquakes generally have a larger Ms magnitude than explosions. This paper proposes a hypothesis test (screening analysis) using Ms and mb that expressly accounts for physical correction model inadequacy in the standard error of the test statistic. With this hypothesis test formulation, the 2009 Democratic People's Republic of Korea (DPRK) announced nuclear weapon test fails to reject the null hypothesis H0: Explosion Characteristics.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2011		**00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE	
Sources of Error and the Statistical Formulation of Ms | mb Seismic Event Screening Analysis for the Comprehensive Nuclear-Test-Ban Treaty (CTBT)	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Los Alamos National Laboratory,P.O. Box 1663 ,Los Alamos,NM,87545**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Published in the Proceedings of the 2011 Monitoring Research Review - Ground-Based Nuclear Explosion Monitoring Technologies, 13-15 September 2011, Tucson, AZ. Volume I. Sponsored by the Air Force Research Laboratory (AFRL) and the National Nuclear Security Administration (NNSA). U.S. Government or Federal Rights License

14. ABSTRACT

The Comprehensive Nuclear-Test-Ban Treaty (CTBT), a global ban on nuclear explosions, is currently pre-entry into force. Under the CTBT, a monitoring system of seismic, hydroacoustic, infrasonic and radionuclide sensors operates and data from this system is analyzed by the International Data Centre (IDC). The IDC provides CTBT signatories basic seismic event parameters and a screening analysis indicating whether an event exhibits explosion characteristics (for example, shallow depth). An important component of the screening analysis is a statistical test of the the null hypothesis H0: Explosion Characteristics using empirical measurements of seismic energy (magnitudes). Relative to mb, earthquakes generally have a larger Ms magnitude than explosions. This paper proposes a hypothesis test (screening analysis) using Ms and mb that expressly accounts for physical correction model inadequacy in the standard error of the test statistic. With this hypothesis test formulation, the 2009 Democratic People's Republic of Korea (DPRK) announced nuclear weapon test fails to reject the null hypothesis H0: Explosion Characteristics.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

OBJECTIVES

Network averaging cannot overcome bias in a magnitude correction model. A new formulation of the CTBT Ms | mb seismic event screening hypothesis test is developed and demonstrated in this paper (Ms | mb denotes the conditional probability model Ms given mb). The formulation properly partitions total error into Model Error and Station Noise, and this partition provides for the correct reduction of the standard error of the hypothesis test. With the correct hypothesis test formulation, a decision of fail to reject the null hypothesis of Explosion Characteristics (denoted H0 or H0: Explosion Characteristics) is made with a network of 27 global International Seismological Centre (ISC) seismic stations for the 2009 DPRK nuclear weapon test. Without the proper error partition in the standard error, H0 is rejected and the 2009 DPRK nuclear weapon test is screened out.

RESEARCH ACCOMPLISHED

The CTBT, in basic obligation, is simple and direct. Article I reads:

- Each State Party undertakes not to carry out any nuclear weapon test explosion or any other nuclear explosion, and to prohibit and prevent any such nuclear explosion at any place under its jurisdiction or control.
- Each State Party undertakes, furthermore, to refrain from causing, encouraging, or in any way participating in the carrying out of any nuclear weapon test explosion or any other nuclear explosion.

Protocol for the CTBT calls for the implementation of a monitoring system and associated IDC with the responsibility to “receive, collect, process, analyse, report on and archive data from monitoring system facilities.” Protocol also directs the IDC to provide an event screening service to the treaty signatories, the technical details of which are specified in the CTBT operational manual.

Seismic Event Screening for the International Data Centre

Seismic energy is generated by earthquakes, volcanoes, mining and oil exploration explosions, natural fracturing of large rocks, large above-ground explosions and nuclear weapon tests. A seismic waveform is a measured transient time series of this energy with distinct segments (phases). Surface wave energy relative to the energy in the initial P-phase, is the basis for an IDC event screening hypothesis test (Ms | mb). The path and distance between event and stations is different and if the phase energy measurements from each station could be accurately corrected for path effects, the measurements would represent energy at the source. Station magnitudes are averaged to estimate an event (network) magnitude. Many surface wave corrections and scales have been developed during the past century. Most notable are scales with distance corrections (Gutenberg [1945], Vaněk et al. [1962], von Seggern [1977], Herak and Herak [1993], Rezapour and Pearce [1998], Stevens and McLaughlin [2001], Bormann et al. [2009]), and scales with corrections for filter effect, distance and path (Marshall and Basham [1972], Russell [2006]). With calibration analysis, the hypothesis formulation developed in this paper is applicable to all commonly used Ms calculations. In general, the model for computed station Ms is

$$\text{Ms} = \log10(\text{Amplitude}) + \text{Path} + \text{Distance} + \text{Filter Effect}. \qquad (1)$$

The Rayleigh wave magnitude Ms is further modeled as a function of mb giving the conditional station magnitude Ms | mb. The IDC provides CTBT signatories basic seismic event parameters including event location and depth, measures of event size (magnitudes), and a screening analysis for events with mb greater than 3.5.

The physical basis of the Ms | mb discriminant is quite mature (see Douglas et al. [1971] and Stevens and Day [1985]), and is based on the physics that for a given mb, a shallow earthquake excites relatively more surface-wave energy than a single-point explosion. However, deep earthquakes like explosions have small Ms for their mb. This means that single-point underground explosions and deep earthquakes will usually fail to reject the null hypothesis of explosion characteristics in the Ms | mb screening analysis. Erroneously screening out a nuclear explosion in IDC analysis is clearly serious and the ramifications of such an error could be political, financial, environmental – or most serious, the loss of life though military provocation. IDC seismic event screening includes a statistical test of the null hypothesis that a seismic event has explosion characteristics, with a very small probability of incorrectly rejecting this hypothesis. This conservative IDC event screening analysis retains events of no concern as well as explosions in a “fail to reject” bulletin provided to CTBT signatories. A technical review of general seismic monitoring is provided in Anderson et al. (2010).

Statistical Model

The probability model of Ms corrected for mb (Ms | mb) is

$$Y = Ms - \eta(mb) = \mu + \text{Model Error} + \text{Station Noise} \quad (2)$$

where η(mb) is a model of the magnitude near the source and Station Noise represents measurement and ambient noise with zero mean. Note that Equation (2) is a regression model of Ms versus mb embedded into a simple random effects analysis of variance model. Model Error is a zero mean random effect that varies from event to event and represents correction model inadequacy from local effects such as inaccurate depth and local material properties, filter/path/distance corrections, and magnitude corrections. These effects are in fact physical and deterministic, yet realistically unknown. The technical approach in Equation (2) is to model these effects as random and properly include the Model Error variance component in calculations of the standard error for the hypothesis test. The variance component for Model Error decreases with improved corrections and physical theory, and the term for Station Noise in the standard error is reduced through station averaging. Importantly, station averaging cannot reduce Model Error. For IDC screening analysis a simple linear regression formulation, $\eta(mb) = \beta \times mb$, is used for both the null and alternate hypothesis. The variance component for Model Error is equal across both the null and alternate hypotheses as is the variance component for Station Noise. The differences in the models for the two hypotheses is represented through differences in the null and alternate population means μ. The the hypothesis test constructed from Equation 2 is composed of a more sophisticated and correct standard error that includes the variance components for Model Error and Station Noise.

Demonstration Analysis

The model component $\eta(mb) = \beta \times mb$ is a physical correction and is assumed known (for IDC event screening β is known). Selby et al. (2011) determine that $\beta = 1$ adequately represents the general physical scaling relationship between Ms versus mb, so that corrected-for-magnitude network surface energy is Ms − mb. The null and alternate hypothesis population means, and variance components for Model Error and Station Noise, are assumed known through calibration analysis demonstrated in the following sections, and the calibrated screening hypothesis test is applied to the 2009 DPRK nuclear weapon test (NWT). The 2009 DPRK NWT was not included in the calibration analysis.

The Ms given mb discriminant is demonstrated with seismic event data acquired from the ISC and the AWE Blacknest Seismological Centre (BSC). The events acquired from the ISC and BSC spanned 1964 to 2000, and included USSR, Chinese, and French underground nuclear explosion tests – 59 explosions for the H0 population; and earthquakes and mining activity – 129 events for the HA population. The ISC and BSC event catalogues provide the mb and the individual station Ms values for each event.

Bootstrap calibration analysis assumes that the calibration data set is a representative sample although possibly not large enough to adequately represent extremes necessary to confidently estimate the variance components. Calibration analysis is accomplished with bootstrap sampling of corrected Ms values ($Ms - \beta \times mb$), and the associated event and station indices. The calibration data included 59 + 129 = 188 calibration events each observed by varying numbers and locations of stations. For example, one event had 56 stations observing, and another had 50 stations observing. Also, 12 events had 3 stations observing and 50 events had two stations observing (at least two stations were required for an Ms calculation in the analysis). All events with stations observing gave 1906 total event/station records, and so each bootstrap sample had 1906 randomly selected records. A total of 5000 bootstrap samples were taken in the calibration analysis. For each bootstrap sample, the null and alternate means μ_0 and μ_A were computed. Also, for each bootstrap sample, the variance components for the one-way random effects model were computed. This collection of 5000 bootstrap variance components provides a technically reasonable approach to inflate the calibration variance components to more conservative values (e.g., the 95th quartile).

Screening Analysis on the 2009 DPRK Nuclear Weapon Test

The 2006 and 2009 DPRK announced NWTs had large network Ms relative to mb (see Bonner et al. [2008] and Patton and Taylor [2008] for research on the 2006 DPRK NWT). This resulted in reconsidering event screening processes at the IDC which led to the recommendation by Selby et al. (2011) of $\beta = 1$. The development in this paper builds on that finding. The ISC/BSC data used in this analysis did not have Ms measurements for the 2006 DPRK NWT. For the 2009 DPRK NWT, twenty seven Ms measurements were reported in the ISC bulletin and the

event mb magnitude was 4.62. Application of the proposed hypothesis test formation in this paper gave a “fail to reject H0” p − value = 0.15. The same analysis with the hypothesis formulation with Model Error rejects H0 with a p − value approximately equal to zero – the wrong decision.

CONCLUSIONS AND RECOMMENDATIONS

Network averaging cannot reduce bias in an magnitude correction model. A new formulation of CTBT seismic event screening hypothesis test is developed and demonstrated in this paper. The formulation properly partitions total error into Model Error and Station Noise, and this partition provides for the correct reduction of the standard error of the hypothesis test. With the correct hypothesis test formulation, a decision of “fail to reject H0: Explosion Characteristics” is made with a network of 27 global ISC seismic stations for the 2009 DPRK nuclear weapon test. Without the proper error partition in the standard error, H0 is rejected and the 2009 DPRK nuclear weapon test is screened out.

ACKNOWLEDGEMENTS

We thank Dmitry Storchak, Director of the International Seismological Centre for his support in the acquisition of the data used in this article.

REFERENCES

Anderson, D. N., G. E. Randall, R. W. Whitaker, S. J. Arrowsmith, M. D. Arrowsmith, D. K. Fagan, N. D. Taylor, S. R.and Selby, F. R. Shult, G. D. Kraft, and W. R. Walter (2010). *Seismic Event Identification*. Wiley Interdisciplinary Reviews: Computational Statistics 2, 414–432.

Anderson, D. N., R. H. Shumway, R. R. Blandford, and S. R. Taylor (2010). Statistical methods in seismology. *Wiley Interdisciplinary Reviews: Computational Statistics* 2: 303–316.

Anderson, D. N., W. R. Walter, D. K. Fagan, T. M. Mercier, and S. R. Taylor (2009). Regional multi-station discriminants: Magnitude, distance and amplitude corrections and sources of error. *Bull. Seismol. Soc. Am*. 99: 794–808.

Bolt, B. (1976). Nuclear explosions and earthquakes: *The Parted Veil*. San Francisco: W. H. Freeman and Company.

Bonner, J. L., R. B. Herrmann, and M. Harkrider, D. G. Pasyanos (2008). The surface wave magnitude for the 9 October 2006 North Korean nuclear explosion. *Bull. Seismol. Soc. Am*. 98: 2598–2506.

Bormann, P., R. Liu, Z. Xu, K. Ren, L. Zhang, and S.Wendt (2009). First application of the new IASPEI teleseismic magnitude standards to data of the China national seismograph network. *Bull. Seismol. Soc. Am*. 99: 1868–1891.

Dahlman, O. and H. Israelson (1977). *Monitoring Underground Nuclear Explosions*. Amsterdam: Elsevier-North Holland.

Douglas, A., J. A. Hudson, and V. K. Kembhavi (1971). The relative excitation of body and surface waves by point sources. *Geophys. J. Roy. Astr. Soc*. 23: 451–460.

Gutenberg, P. (1945). Amplitudes of surface waves and the magnitudes of shallow earthquakes. *Bull. Seismol. Soc. Am*. 35: 3–12.

Herak, M. and D. Herak (1993). Distance dependence of Ms and calibrating function for 20 second Rayleigh waves. *Bull. Seismol. Soc. Am*. 83: 1881–1892.

Marshall, P. D. and P. Basham (1972). Discrimination between earthquakes and underground explosions employing an improved Ms scale. *Geophys. J. R. Astr. Soc*. 29: 431–458.

Mikhailov, V. N. (1999). *Catalog of Worldwide Nuclear Testing*. New York: Begell-Atom.

Patton, H. J. and S. R. Taylor (2008). Effects of shock-induced tensile failure on mb –Ms discrimination: Contrasts between historic nuclear explosions and the North Korean test of 9 October 2006. *Geophys. Res. Lett*. 35: L14301.

Rezapour, M. and R. G. Pearce (1998). Bias in surface-wave magnitude Ms due to inadequate distance correction. *Bull. Seismol. Soc. Am*. 88: 43–61.

Russell, D. R. (2006). Development of a time-domain, variable-period surface-wave magnitude measurement procedure for application at regional and teleseismic distances, Part I: Theory. *Bull. Seismol. Soc. Am*. 96: 665–677.

Scheffe, H. (1959). *The Analysis of Variance*. New York: John Wiley & Sons.

Searle, S. R. (1971). *Linear Models*. New York: John Wiley & Sons.

Searle, S. R., G. Casella, and C. E. McCulloch (1992). Variance Components. New York: John Wiley & Sons.

Selby, N. D., P. D. Marshall, and D. Bowers (2011). mb : Ms event screening revisited. *Bull. Seismol. Soc. Am*, in review.

Stevens, J. L. and S. M. Day (1985). The physical basis of the mb : Ms and variable frequency magnitude methods for earthquake/explosion discrimination. *J. Geophys. Res*. 90: 3009–3020.

Stevens, J. L. and K. L. McLaughlin (2001). Optimization of surface wave identification and measurement. *Pure. App. Geophys*. 158, 1547–1582.

Vaněk, J., A. Zatopek, V. Karnik, Y. Riznichenko, E. Saverensky, S. Solov'ev, and N. Shebalin (1962). Standardization of magnitude scales. Bull. (Izvest.) Acad. Sci. U.S.S.R., Geophys. Ser. 2, 153–158.

von Seggern, D. (1977). Amplitude distance relation for 20-second Rayleigh waves. *Bull. Seismol. Soc. Am*. 67, 405–411.